Optimization Methods and Software, 2001, vol. 16, pp. 231–256
Reprints available directly from the publisher
Photocopying permitted by license only

GENERAL LAGRANGE-TYPE FUNCTIONS IN CONSTRAINED GLOBAL OPTIMIZATION PART II: EXACT AUXILIARY FUNCTIONS

[a] Yu.G. EVTUSHENKO[1], [b]A.M. RUBINOV and [a]V.G. ZHADAN

[a] Computing Centre of Russian Academy of Sciences,
40 Vavilov Str., 117967 Moscow GSP-1, Russia;
[b] School of Information Technology and Mathematical Sciences
University of Ballarat, Victoria 3353, Australia

This paper is a continuation of [13]. For each constrained optimization problem we consider certain unconstrained problems, which are constructed by means of auxiliary (Lagrange-type) functions. We study only exact auxiliary functions, it means that the set of their global minimizers coincides with the solution set of the primal constrained optimization problem. Sufficient conditions for the exactness of an auxiliary function are given. These conditions are obtained without assumption that the Lagrange function has a saddle point. Some examples of exact auxiliary functions are given.

Keywords: Constrained optimization; Auxiliary function; Exact auxiliary function; Convolution function; Generalized polar function

1 INTRODUCTION

This paper is the second part of [13]. We consider the nonlinear programming problem:

$$P(f,g) : f(x) \to \min \quad \text{subject to} \quad x \in X, \quad g(x) \le 0,$$

where X is a metric space, $f(x)$ and $g(x) = (g^1(x), \dots, g^m(x))$ are functions mapping X into the real line $\mathbb{R}$ and into the m-dimensional Euclidean space $\mathbb{R}^m$, respectively. We assume that the set of feasible elements $X_0 = \{x \in X : g(x) \le 0\}$ is not empty and the problem $P(f,g)$ has a solution. Let ρ be the optimal value of $P(f,g)$:

$$\rho = \min\{f(x) : x \in X_0\}, \tag{1}$$

and let $X_* = \{x \in X_0 : f(x) = \rho\}$ be the solution set.

Various auxiliary functions for problem $P(f,g)$, which are similar in certain sense to the classical Lagrange function were introduced in [13]. These functions are defined on $X \times \Omega$, where Ω is a set of parameters. Let $M(x,\omega)$ be an auxiliary function, $X_0 \subseteq Y \subseteq X$ and $\omega_* \in \Omega$. We will examine an "unconstrained" problem of the form:

$$M(x,\omega_*) \to \min \quad \text{subject to} \quad x \in Y. \tag{2}$$

Here we study auxiliary functions $M(x,\omega)$ which enjoy the following property: there exists a set $\Omega_* \subseteq \Omega$ such that for any $\omega_* \in \Omega_*$ the solution set of (2) coincides with the solution set

[1]Corresponding author. Fax: 7-095 135 6159. E-mail: evt@ccas.ru

X_* of $P(f,g)$. Such auxiliary functions are called *exact auxiliary* (EA) functions. We survey known results related to EA functions and present some new results.

The well-known representatives of EA functions are exact penalty (EP) functions introduced and studied by I.I. Eremin [8] and W. Zangwill [27] in 1967. Later on these functions were investigated by many authors. We mention only [1, 2, 4, 6, 7, 9, 14, 18, 21, 22, 26]. EP functions are very important from theoretical and practical point of view. It turned out that during a long period of time other EA functions were not considered. More than a decade ago it became clear that among auxiliary functions, which are used in various realizations of the method of centers there are also EA functions. The estimation (either upper or lower) of the optimal value of $P(f,g)$ appears as a parameter for this kind of EA functions. This estimation is used instead of penalty coefficient to provide exactness. The theoretical study of such type auxiliary functions were carried out in papers [10, 11], where the notion of EA function was first introduced. Several other examples of EA functions were proposed in these papers as well.

As in [13] we are mainly concerned with auxiliary functions of the form

$$M(x,\eta) = \varphi(f(x) - \eta, \psi(g(x))), \tag{3}$$

where φ and ψ are convolution functions (see, for detail, [13]), and η is an estimation of the optimal value ρ. We also shall consider convolution functions

$$M(x,\eta,\omega) = \varphi(f(x) - \eta, \psi(g(x);\omega)) \tag{4}$$

depending on a parameter $\omega \in \Omega$. Necessary and sufficient optimality conditions for $P(f,g)$ were obtained in [13] in terms of functions (4). A penalty function can be presented in the form (4) if the outer convolution function φ is linear and the function ψ possesses some natural properties. If φ differs from a linear function we obtain in particular various auxiliary functions, which were used in the study of methods of centers (external or internal).

The method of centers was introduced by P. Huard [15] in 1964. A suitably defined auxiliary function was introduced and applied in this paper. This function is convex in the interior of feasible set, and is infinite on its boundary, so it is attained its minimum at an interior point. Unconstrained minima of this function converge to the solution of the constrained problem. The method of external centers was proposed by D. Morrison [17]. All minimizers of the auxiliary function, which is exploited in this method, do not belong to the feasible region. Moreover, a lower estimation of the optimal value is used, which is updated at each step of the iterative process. Some other variants of the method of centers were considered in [25, 5, 16].

The property for auxiliary function being exact depends both on the structure of the problem $P(f,g)$ and on the type of convolution functions used in (3) or (4). Our aim here is to show that this property is not inherent only to penalty functions for which φ is a linear convolution function. It turns out that, if the convolution function ψ has the form, similar to those used in EP functions, the auxiliary functions (3) or (4) with the nonlinear function φ may be also exact.

In [10, 11] all results related to EA functions were obtained for a problem $P(f,g)$ such that its classical Lagrange function has a saddle point. In this paper we examine more general class of problems. We assume only that a certain generalization of the Lagrange function has a saddle point. It allows one to consider much more broad class of problems. We express conditions providing exactness in terms of generalized polar functions.

The paper is organized as follows. In Section 2 we formulate sufficient optimality conditions based on the function (4) where the outer convolution φ is linear. In Section 3 we introduce the notion of generalized polar function and give some examples of such functions. Various classes of exact auxiliary functions are presented in Section 4.

Since this paper is a continuation of [13], we use terminology and notations from [13] without special explanations.

2 AUXILIARY FUNCTIONS WITH LINEAR OUTER CONVOLUTION FUNCTION

The aim of this section is to make more precise the results obtained in [13] for linear outer convolution functions φ in (4). These results, relating mainly to sufficient conditions for optimality, will be used in the study of the problem $P(f,g)$.

Let $Y \in \mathcal{Y} := \{Y : X_0 \subseteq Y \subseteq X\}$. Denote by $\mathcal{T}(Y)$ the image of Y under the mapping $(f,g) : Y \to \mathbb{R}^{1+m}$, that is, the set

$$\mathcal{T}(Y) = \{(u,v) \in \mathbb{R}^{1+m} : u = f(x),\ v = g(x),\ x \in Y\}.$$

Denote also by $\mathcal{T}_\eta(Y)$ the η-shift of $\mathcal{T}(Y)$ along the first coordinate: $\mathcal{T}_\eta(Y) = \mathcal{T}(Y) - \eta e_0$, where $e_0 = (1, 0, \ldots, 0)$ is the unit vector. Recall (see [13]) that optimality conditions, which we consider, express the separability of the set $\mathcal{T}_\rho(Y)$ with ρ defined by (1) and the set

$$\mathcal{H}^- = \{(u,v) \in \mathbb{R}^{1+m} : u < 0,\ v \leq 0\}$$

by a non-necessarily linear function.

Consider the outer convolution function $\varphi(u,w) = u + w$ and assume that an inner convolution function ψ depends on a parameter $\omega \in \Omega$. Due to a special form of φ we can consider the auxiliary function (4) without the parameter η (see [13] for details). Then the auxiliary function $M \equiv M_\psi$ has a form:

$$M_\psi(x,\omega) = f(x) + \psi(g(x);\omega)), \qquad (x \in X,\ \omega \in \Omega). \tag{5}$$

Let $\psi(\cdot;\omega)$ be a proper convolution function for each $\omega \in \Omega$ and let M_ψ be the corresponding auxiliary function (5).

Definition 2.1. *We say that the triplet (ψ, M_ψ, Y) enjoys the property* **(A)** *if there exist $x_* \in X_0$ and $\omega_* \in \Omega$ such that*

$$M_\psi(x_*,\omega_*) = \min_{x \in Y} M_\psi(x,\omega_*), \tag{6}$$

$$\psi(g(x_*);\omega_*) = 0. \tag{7}$$

The following assertion holds.

Proposition 2.1. *Let $Y \in \mathcal{Y}$. Let $\psi(\cdot;\omega)$ be a proper convolution function for any $\omega \in \Omega$. If (ψ, M_ψ, Y) satisfies the property* **(A)***, then x_* is a solution of $P(f,g)$.*

Proof. It follows from [13, Proposition 6.1]. □

Remark 2.1. *Let a triplet (ψ, M_ψ, Y) enjoy the property* **(A)** *and let (x_*,ω_*) be a point such that (6) and (7) hold. Since $\psi(\cdot;\omega)$ is a proper convolution function and $g(x_*) \leq 0$, it follows that*

$$\psi(g(x_*);\omega_*) = 0 \geq \psi(g(x_*);\omega) \ \ \textit{for all} \ \ \omega \in \Omega,$$

hence,

$$M_\psi(x_*,\omega_*) = f(x_*) \geq M_\psi(x_*,\omega) \ \ \textit{for all} \ \ \omega \in \Omega.$$

On the other hand, (6) *implies*

$$M_\psi(x_*, \omega_*) \le M_\psi(x, \omega_*) \quad \text{for all} \quad x \in Y.$$

Thus, (x_*, ω_*) *is a saddle point of the function* $M(x, \omega)$ *on the set* $Y \times \Omega$*, that is,*

$$M_\psi(x, \omega_*) \ge M_\psi(x_*, \omega_*) \ge M_\psi(x_*, \omega) \quad \forall\, x \in Y \quad \text{and} \quad \forall\, \omega \in \Omega. \tag{8}$$

Assume now that the function ψ *possesses the following property:*

$$\sup\{\psi(v; \omega) : \omega \in \Omega\} = 0 \quad \text{for all} \quad v \in \mathbb{R}^m_-. \tag{9}$$

Then (8) *implies both* (6) *and* (7). *Indeed, it follows directly from* (8) *that* $M_\psi(x_*, \omega_*) = \min_{x \in Y} M_\psi(x, \omega_*)$. *We also have*

$$\psi(g(x_*); \omega_*) = \sup_{\omega \in \Omega} \psi(g(x_*); \omega) = 0.$$

Thus, if (9) *holds, then the property* **(A)** *is equivalent to the existence of a saddle point of the function* M_ψ *on the set* $Y \times \Omega$*.*

For a function $\psi : \mathbb{R}^m \times \Omega \to \mathbb{R}$ consider the set

$$\mathcal{L}^+_\psi(\omega) = \{(u, v) \in \mathbb{R}^{1+m} : h(u, v; \omega) \ge 0\},$$

where $h(u, v; \omega) = u + \psi(v; \omega)$ is the separation function generated by the function $\phi(u, w) = u + w$ and the function $\psi(v; \omega)$ (see [13] for details).

Proposition 2.2. *Assume that* $\psi(\cdot; \omega)$ *is a proper convolution function for each* $\omega \in \Omega$*; assume also that* (9) *holds. Then* $(x_*, \omega_*) \in X_0 \times \Omega$ *is a saddle point of the function* $M_\psi(x, \omega)$ *on* $Y \times \Omega$ *if and only if*

$$\mathcal{T}_\eta(Y) \subseteq \mathcal{L}^+_\psi(\omega_*), \quad \text{where} \quad \eta = f(x_*). \tag{10}$$

Proof. Let (10) hold. Then

$$f(x) - \eta + \psi(g(x); \omega_*) \ge 0, \qquad (x \in Y). \tag{11}$$

Since $x_* \in X_0$, we have

$$\psi(g(x_*); \omega_*) \le 0. \tag{12}$$

Combining (11) and (12), we conclude that

$$M_\psi(x, \omega_*) = f(x) + \psi(g(x); \omega_*) \ge \eta = f(x_*) \ge f(x_*) + \psi(g(x_*); \omega_*) = M_\psi(x_*, \omega_*).$$

It follows from (11) (with $x = x_*$) that $\psi(g(x_*); \omega_*) \ge 0$, so due to (12) we have $\psi(g(x_*); \omega_*) = 0$, hence

$$\psi(g(x_*); \omega_*) = 0 = \max\{\psi(g(x_*); \omega) : \omega \in \Omega\}.$$

Thus, (x_*, ω_*) is a saddle point of the function M_ψ on $Y \times \Omega$.

Assume now that (x_*, ω_*) is a saddle point. Then we have

$$f(x_*) + \psi(g(x_*); \omega_*) = M_\psi(x_*, \omega_*) = \max_{\omega \in \Omega} M_\psi(x_*, \omega) = f(x_*) + \max_{\omega \in \Omega} \psi(g(x_*); \omega).$$

Since (9) holds, it follows that

$$\psi(g(x_*); \omega_*) = \max_{\omega \in \Omega} \psi(g(x_*), \omega) = 0. \tag{13}$$

Let $x \in Y$. Applying (13), we conclude that

$$f(x) + \psi(g(x);\omega_*) = M_\psi(x,\omega_*) \geq M_\psi(x_*,\omega_*) = f(x_*) + \psi(g(x_*);\omega_*) = \eta + \psi(g(x_*);\omega_*) = \eta.$$

Thus, $f(x) - \eta + \psi(g(x);\omega_*) \geq 0$ for all $x \in Y$, so $\mathcal{T}_\eta(Y) \subseteq \mathcal{L}^+_\psi(\omega_*)$. □

Corollary 2.1. *Let* (10) *hold for some* (x_*, ω_*). *Then* $M_\psi(x,\omega)$ *with* ψ *satisfying assumptions of Proposition* 2.2 *has a saddle point.*

The simplest example of a proper convolution function ψ is a linear function

$$\psi(v;\omega) = \langle \omega, v\rangle, \qquad \omega \in \Omega := \mathbb{R}^m_+, \tag{14}$$

where $\langle \cdot,\cdot\rangle$ stands for the inner product. Clearly, this function satisfies (9), therefore, the property (A) is equivalent to the existence of a saddle point of M_ψ on $Y \times \mathbb{R}^m_+$. The function M_ψ with ψ of (14) coincides with the classical Lagrange function.

It follows from Proposition 2.2 that the existence of a saddle point (x_*, ω_*) of the Lagrange function is equivalent to the following assertion: $\mathcal{T}_\rho(Y)$ belongs to half-space $\mathcal{L}^+_\psi(\omega_*)$, where $\mathcal{L}^+_\psi(\omega)$ is defined by

$$\mathcal{L}^+_\psi(\omega) = \{(u,v) \in \mathbb{R}^{1+m} : \langle (u,v),(1,\omega)\rangle \geq 0\}.$$

We now give two more examples:

Example 2.1. Consider the function

$$\psi(v;\omega) = \max[\langle \omega_1, v\rangle, \langle \omega_2, v\rangle], \qquad \omega = [\omega_1,\omega_2] \in \Omega := \mathbb{R}^m_+ \times \mathbb{R}^m_+; \tag{15}$$

It is easy to check that $\psi(\cdot;\omega)$ is a proper convolution function for any $\omega \in \Omega$.

Example 2.2. Consider the function

$$\psi(v;\omega) = \frac{\alpha}{2}\left[\left\| \left(\frac{v}{\alpha} + \omega_1\right)_{[0,\omega_1+\omega_2/\alpha]}\right\|^2 - \|\omega_1\|^2\right], \tag{16}$$

where

$$\omega = [\omega_1,\omega_2,\alpha] \in \Omega := \mathbb{R}^m_+ \times \mathbb{R}^m_{++} \times \mathbb{R}_{++},$$

and $(x)_{[a,b]}$ is the orthogonal projection of a vector x onto the box set $\{x \in \mathbb{R}^m : a^i \leq x^i \leq b^i,$ $1 \leq i \leq m\}$. It can be shown that $\psi(\cdot;\omega)$ is a proper convolution function for any $\omega \in \Omega$ (see [13] for details).

Both functions (15) and (16) were introduced in [12]. In the limit case, when $\omega_2^i = +\infty$, $1 \leq i \leq m$, the function (16) transfers to the augmented Lagrangian [3].

For many problems $P(f,g)$ it is possible to find a proper convolution function ψ and a set $Y \in \mathcal{Y}$ such that the triplet (ψ, M_ψ, Y) enjoys property **(A)**, even the classical Lagrange function for this problem has no saddle points.

Consider an inner convolution function ψ defined by (15) from this point of view. Exploiting an outer convolution function $\varphi(u,w) = u + w$ and the function ψ, we obtain the separation function

$$h(u,v;\omega) = u + \max[\langle \omega_1, v\rangle, \langle \omega_2, v\rangle]. \tag{17}$$

The auxiliary function M_ψ, corresponding to this separation function (see [13, Section 6]), has the form

$$M_\psi(x,\omega) = f(x) + \max[\langle \omega_1, g(x)\rangle, \langle \omega_2, g(x)\rangle].$$

Clearly, ψ possesses (9), so the property **(A)** for the triplet (ψ, M_ψ, Y) is equivalent to the existence of a saddle point of the function M_ψ on $Y \times \Omega$. It follows from Proposition 2.2 that a saddle point exists if and only if $\mathcal{T}_\rho(Y) \subseteq \mathcal{L}_\psi^+(\omega_*)$ for some $\omega_* \in \Omega$, where

$$\mathcal{L}_\psi^+(\omega) = \{(u,v) \in \mathbb{R}^{1+m} : u + \max[\langle \omega_1, v\rangle, \langle \omega_2, v\rangle] \geq 0\}.$$

Let

$$\mathcal{L}_i^+(\omega) = \{(u,v) \in \mathbb{R}^{1+m} : \langle (u,v), (1,\omega_i)\rangle \geq 0\}, \qquad i = 1,2.$$

Since $u + \max_{i=1,2} \langle \omega_i, v\rangle = \max_{i=1,2}(u + \langle \omega_i, v\rangle)$, it follows that $\mathcal{L}_\psi^+(\omega)$ can be presented as the union of half-spaces $\mathcal{L}_1^+(\omega)$ and $\mathcal{L}_2^+(\omega)$.

Thus, property (A) holds for the triplet (ψ, M_ψ, Y) with ψ of (15) if and only if $\mathcal{T}_\rho(Y)$ is contained into the union of two half-spaces, which are defined by vectors $(1,(\omega_*)_1)$ and $(1,(\omega_*)_2)$, respectively, where $(\omega_*)_i \in \mathbb{R}_+^m$, $i = 1,2$. (Recall that property (A) holds for (ψ, M_ψ, Y), where M_ψ is the classical Lagrange function, if and only if $\mathcal{T}_\rho(Y)$ is a subset of a half-space defined by a vector $(1,\omega_*)$).

It is clear that (17) is an increasing positively homogeneous (IPH) function on $\mathbb{R}^{1+m}$. Moreover, if $\omega_1 \neq 0$ and $\omega_2 \neq 0$, then $h(u,v;\omega)$ is a regular weak separation function (see [13, Section 4]).

The function

$$\psi(v;\omega) = \max_{i=0,1,\dots,p} \langle \omega_i, v\rangle, \tag{18}$$

where

$$\omega = (\omega_0, \omega_1, \dots, \omega_p), \quad \omega_i \in \mathbb{R}_+^m, \quad i = 0,1,\dots,p, \quad p > 2,$$

is a natural generalization of (15). The corresponding separation function and auxiliary function have the following forms respectively:

$$h(u,v;\omega) = u + \max[\langle \omega_0, v\rangle, \langle \omega_1, v\rangle, \dots, \langle \omega_p, v\rangle], \tag{19}$$

$$M_\psi(x,\omega) = f(x) + \max[\langle \omega_0, g(x)\rangle, \langle \omega_1, g(x)\rangle, \dots, \langle \omega_p, g(x)\rangle]. \tag{20}$$

Here $\omega \in \Omega := \mathbb{R}_+^{m(1+p)}$. Note that the function ψ defined by (18) enjoys (9).

Assume that there exist $x_* \in X_0$ and $\omega_* \in \Omega$ such that (6) and (7) hold. Then, applying Proposition 2.1, we conclude that x_* is a solution of $P(f,g)$. Consider the set $\mathcal{L}_\psi^+(\omega)$ corresponding to the separation function (19). Let $J_p = \{0,1,\dots,p\}$. Then

$$\mathcal{L}_\psi^+(\omega) = \{(u,v) : h(u,v;\omega) \geq 0\} = \{(u,v) : (\exists\, i \in J_p) u + \langle \omega_i, v\rangle \geq 0\},$$

so

$$\mathcal{L}_\psi^+(\omega) = \bigcup_{i \in J_p} \{(u,v) : u + \langle \omega_i, v\rangle \geq 0\}.$$

Thus, $\mathcal{L}_\psi^+(\omega)$ can be represented as the union of $(p+1)$ half-spaces

$$\mathcal{L}_i^+(\omega) = \{(u,v) \in \mathbb{R}^{1+m} : \langle (u,v), (1,\omega_i)\rangle \geq 0\}, \qquad i \in J_p,$$

which are defined by vectors $(1,\omega_i)$ with $\omega_i \in \mathbb{R}_+^m$, $(i \in J_p)$. It follows from Proposition 2.2 that the auxiliary function M_ψ has a saddle point on $Y \times \Omega$ if and only if a vector $\omega_* = ((\omega_*)_i)_{i=0}^p \geq 0$ can be found such that $\mathcal{T}_\rho(Y) \subset \bigcup_{i \in J_p} \mathcal{L}_i^+(\omega_*)$, where $\rho = \min_{x \in X_0} f(x)$ is the value of the problem $P(f,g)$.

Consider now the complement $\mathcal{L}_\psi^-(\omega)$ to the set $\mathcal{L}_\psi^+(\omega)$. Clearly,

$$\mathcal{L}_\psi^-(\omega) = \{(u,v) : u + \langle \omega_i, v\rangle < 0, \quad i \in J_p\} \tag{21}$$

is an open convex cone. Let $(u,v) \in \mathcal{H}^-$, that is, $u < 0$, $v \le 0$. Since $\omega_i \ge 0$, it follows that $u + \langle \omega_i, v\rangle < 0$, so $\mathcal{H}^- \subset \mathcal{L}_\psi^-(\omega)$. We have proved the following assertion:

Proposition 2.3. *Let ψ be a convolution function defined by* (18) *and $Y \in \mathcal{Y}$. Then an auxiliary function* (20) *has a saddle point on the set $Y \times \Omega$ if and only if there exists a vector $\omega_* = ((\omega_*)_i)_{i\in J_p}$ such that the open convex cone $\mathcal{L}_\psi^-(\omega_*)$ separates the sets $\mathcal{T}_\rho(Y)$ and $\mathcal{H}^-$ in the following sense: $\mathcal{H}^- \subset \mathcal{L}_\psi^-(\omega_*)$ and $\mathcal{T}_\rho(Y) \cap \mathcal{L}_\psi^-(\omega_*) = \emptyset$.*

Remark 2.2. *We can express the statement of Proposition 2.3 in terms of separation function h defined by (19):*

$$\begin{aligned} h(u,v;\omega_*) &< 0 \quad \textit{for all} \quad (u,v) \in \mathcal{H}^-; \\ h(u,v;\omega_*) &\ge 0 \quad \textit{for all} \quad (u,v) \in \mathcal{T}_\rho(Y). \end{aligned}$$

Recall, that $\mathcal{T}_\rho(Y) \cap \mathcal{H}^- = \emptyset$ for any problem $P(f,g)$ and any $Y \in \mathcal{Y}$ (see [13, Proposition 2.1]).

We now show that, if $\mathcal{H}^-$ and $\mathcal{T}_\rho(Y)$ can be separated by a convex cone and $\psi(v;\omega)$ is a convolution function defined by (18) with $p = m$, then auxiliary function M_ψ has a saddle point on the set $Y \times \Omega$.

First we remind the following definition. A convex cone $K \subset \mathbb{R}^n$ is called *simplicial* if K is a convex hull of n linearly independent vectors $z_1, \dots, z_n$. Let K be a simplicial cone. Since $K = \left\{ \sum_{i=1}^n \alpha_i z_i : \alpha_i \ge 0, \ \ i = 1, \dots, n \right\}$, it follows that K is isomorphic to $\mathbb{R}^n_+$. It follows from this isomorphism that there exist n linearly independent vectors $\ell_1, \dots, \ell_n$ such that $K = \{z : \langle \ell_i, z\rangle \le 0 : \ i = 1, \dots, n\}$. In other words,

$$K = \{z \in \mathbb{R}^n : \max_{i=1,\dots,n} \langle \ell_i, z\rangle \le 0\}.$$

Clearly that

$$\operatorname{int} K = \{z \in \mathbb{R}^n : \max_{i=1,\dots,n} \langle \ell_i, z\rangle < 0\},$$

where $\operatorname{int} K$ is an interior of K.

Theorem 2.1. *Consider the problem $P(f,g)$ and a set $Y \in \mathcal{Y}$. Let $\psi(v,\omega)$ be a convolution function defined by* (18) *with $p = m$. Then the auxiliary function M_ψ has a saddle point on the set $Y \times \Omega$ if and only if there exists an open convex cone $\mathcal{K} \subset \mathbb{R}^{1+m}$ which separates $\mathcal{H}^-$ and $\mathcal{T}_\rho(Y)$, that is, $\mathcal{H}^- \subset \mathcal{K}$ and $\mathcal{T}_\rho(Y) \cap \mathcal{K} = \emptyset$.*

Proof.
1. If a saddle point (x_*, ω_*) exists, then the cone $\mathcal{K} = \mathcal{L}_\psi^-(\omega_*)$ defined by (21) satisfies required properties.
2. Assume that an open convex cone $\mathcal{K}$, separating $\mathcal{H}^-$ and $\mathcal{T}_\rho(Y)$ exists. Let $\varepsilon > 0$. Consider unit vectors $\bar{e}_i = (\bar{e}_i^0, \bar{e}_i^1, \dots, \bar{e}_i^m)$, $0 \le i \le m$, with components

$$\bar{e}_i^j = \begin{cases} 1, & j = i, \\ 0, & j \ne i, \end{cases}$$

and vectors $\bar{e}_{i,\varepsilon} = (\bar{e}_{i,\varepsilon}^0, \bar{e}_{i,\varepsilon}^1, \dots, \bar{e}_{i,\varepsilon}^m)$, $0 \le i \le m$, with components

$$\bar{e}_{i,\varepsilon}^j = \begin{cases} 1, & j = i, \\ -\varepsilon, & j \ne i, \end{cases} \qquad 0 \le j \le m.$$

Let

$$s_{i,\varepsilon} = \begin{cases} -\bar{e}_{i,\varepsilon}, & -\bar{e}_i \in \mathcal{K}, \\ -\bar{e}_i, & -\bar{e}_i \notin \mathcal{K}, \end{cases} \qquad 0 \le i \le m,$$

and $\mathcal{K}_\varepsilon$ be a cone hull of vectors $s_{0,\varepsilon}, s_{1,\varepsilon}, \dots, s_{m,\varepsilon}$. Assume that ε is a sufficiently small number. Then $\operatorname{int}\mathcal{K}_\varepsilon \subset \mathcal{K}$ and vectors $s_{i,\varepsilon}$, $(i = 0, \dots, m)$ are linearly independent. Hence, $\mathcal{K}_\varepsilon$ is a simplicial cone. The simpliciality of $\mathcal{K}_\varepsilon$ implies the existence of $m+1$ linear independent vectors $(d_i, \omega_i) \in \mathbb{R}^{1+m}$ such that $\mathcal{K}_\varepsilon = \{(u,v) : d_i u + \langle \omega_i, v\rangle \le 0\}$. It is easy to check that $\mathcal{H}^- \subset \operatorname{int}\mathcal{K}_\varepsilon$. This inclusion implies $\omega_i \ge 0$ and $d_i > 0$ for all $i = 0, 1, \dots, m$. Indeed, we have $d_i u + \langle \omega_i, v\rangle \le 0$ for an arbitrary $v \le 0$ and $u < 0$. Turning u to zero, we have $\langle \omega_i, v\rangle \le 0$ for all $v \le 0$, hence $\omega_i \ge 0$. We also have $\langle (d_i, \omega_i), \bar{e}_0\rangle = d_i \le 0$, so $d_i \ge 0$.

If $d_i = 0$ for an index i, then $\langle (d_i, \omega_i), \bar{e}_0\rangle = 0$. Since

$$\operatorname{int}\mathcal{K}_\varepsilon = \{(u,v) : \langle (d_i, \omega_i), (u,v)\rangle < 0, \quad i = 0, 1, \dots, m\},$$

it follows that $-\bar{e}_0$ is a boundary point of $\mathcal{K}_\varepsilon$. This contradicts inclusions $-\bar{e}_0 \in \mathcal{H}^- \subset \operatorname{int}\mathcal{K}_\varepsilon$. Hence $d_i > 0$. Assume without loss of generality that $d_i = 1$ for all i. Let $\omega_* = (\omega_0, \omega_1, \dots, \omega_m)$. Then

$$\operatorname{int}\mathcal{K}_\varepsilon = \{(u,v) : u + \langle \omega_i, v\rangle < 0, \quad i = 0, 1, \dots, m\} = \mathcal{L}_\psi^-(\omega_*),$$

where $\mathcal{L}_\psi^-(\omega_*)$ is the cone defined by (21). Since $\operatorname{int}\mathcal{K}_\varepsilon \subset \mathcal{K}$ and $\mathcal{K} \cap \mathcal{T}_\rho(Y) = \emptyset$, it follows that also $\operatorname{int}\mathcal{K}_\varepsilon \cap \mathcal{T}_\rho(Y) = \emptyset$. Since $\operatorname{int}\mathcal{K}_\varepsilon \supset \mathcal{H}^-$, we can conclude that $\mathcal{K}_\varepsilon = \mathcal{L}_\psi^-(\omega_*)$ separates $\mathcal{T}_\rho(Y)$ and $\mathcal{H}^-$. The desired result follows now from Proposition 2.3. □

The auxiliary function (20) possesses the remarkable property. Namely, if there exists an open convex cone, which separates $\mathcal{H}^-$ and $\mathcal{T}_\rho(Y)$, then the equality (6) holds not only for the particular value ω_* but also for a fairly broad set of parameters ω. It follows from the fact that we may change e and consequently take various cones $\mathcal{K}_\varepsilon$. Moreover, we may define the cone $\mathcal{K}_\varepsilon$ by means of a vector-parameter $\varepsilon = (\varepsilon^0, \dots, \varepsilon^m)$. This leads to a variety of directional vectors $(d_0, \omega_0), \dots, (d_m, \omega_m)$. Then (6) holds for corresponding vectors ω_* as well. In what follows the functions with such property will be called *exact auxiliary* (EA) functions. These functions will be examined in Section 4.

3 GENERALIZED POLAR FUNCTIONS

A polar function is a very efficient tool for obtaining various evaluations of the image $\mathcal{T}_\rho(Y)$ of the problem $P(f,g)$, especially if assumptions of Proposition 2.1 hold for this problem. The classical notion of polarity closely related to bilinear coupling function. Since we consider a general coupling function $\psi(v;\omega)$ instead of bilinear function $\langle v, \omega\rangle$, we need corresponding generalization of polarity. Various concepts of generalized conjugacy and, in particular, generalized polarity are based on the notion of abstract convexity (see [24] and also [19, 20]). Here we introduce generalized polar functions which are more convenient for our purposes. The proposed definition is based on the Minkowski–Mahler inequality. In other words, we try to define the generalized polar function in order to obtain the best possible inequality of the form $F(z, z_0) \le q(z)q^0(z_0)$ for a coupling function F with a given function q. (Here q^0 is a polar function).

Classical polar functions have been studied only with respect to nonnegative functions q, however, we shall not restrict ourselves by considering only nonnegative functions. We shall consider not only functions defined on the entire space but also its restrictions to some subsets. In such a case we can obtain finite polar functions. In particular, we shall define polar functions to common and interior convolution functions (see [13] for the definition of these functions). We also define and study generalized polarity for the restriction of interior convolution functions to the non-positive orthant.

Let $F : \mathbb{R}^m \times \mathbb{R}^s \to \mathbb{R}$ be a coupling function, and let $Z \subseteq \mathbb{R}^m$. Let further $q : \mathbb{R}^m \to \mathbb{R}$.

For each $z_0 \in \mathbb{R}^s$ consider the set

$$\Lambda(z_0) = \{\lambda \in \bar{\mathbb{R}} : F(z, z_0) \le \lambda q(z) \text{ for all } z \in Z\} = \{\lambda \in \bar{\mathbb{R}} : \Delta_{z_0}(\lambda) \le 0\},$$

where

$$\Delta_{z_0}(\lambda) := \sup_{z \in Z}[F(z, z_0) - \lambda q(z)].$$

Note that $\Lambda(z_0)$ is a closed segment on the extended real line (maybe empty). Let

$$Z_0 = \{z_0 \in \mathbb{R}^s : \Lambda(z_0) \ne \emptyset\}.$$

For each $z_0 \in Z_0$ we are interested in a point $\lambda(z_0)$ from the segment $\Lambda(z_0)$ (which can be equal to $+\infty$ or to $-\infty$), which provides the exact evaluation of F in the following sense: the function $\lambda(z_0)q$ is the least function proportional to q, which is a majorant of $F(\cdot, z_0)$ on Z. It means that $\lambda(z_0)$ is satisfied to the following two conditions:

(1) $\lambda(z_0) \in \Lambda_*(z_0) := \{\lambda_0 \in \Lambda(z_0) : \Delta_{z_0}(\lambda_0) \ge \Delta_{z_0}(\lambda) \;\; \forall\, \lambda \in \Lambda(z_0)\}$;
(2) there does not exist another $\lambda \in \Lambda_*(z_0)$ such that

$$F(z, z_0) \le \lambda q(z) \le \lambda(z_0)q(z) \text{ for all } z \in Z$$

with the right inequality being strict for some $z \in Z$.

Such a point $\lambda(z_0)$ always exists, however, sometimes this point is not unique. Denote by $\Lambda_0(z_0)$ the set of all such $\lambda_0(z_0)$. We now give some examples:

1. Let both $F(z, z_0)$ and $q(z)$ be positive for all $z \in Z$. Then $\Lambda(z_0) \subseteq \mathbb{R}_+$ and $\lambda(z_0)$ is the left end of $\Lambda(z_0)$. The set $\Lambda_0(z_0)$ consists of one point.

2. Let both $F(z, z_0)$ and $q(z)$ be negative for all $z \in Z$. Then $\lambda(z_0)$ is the right end of $\Lambda(z_0)$. The set $\Lambda_0(z_0)$ also consists of one point.

3. Let $Z = [-1, 1]$ and $\Phi(z, z_0) = -0.5 + zz_0$ with $z_0 = 1$. Then for the function $q(z) = \max[0.5z, 2z]$ we have $\Lambda_0(z_0) = \{0.25, 3\}$, that is, $\Lambda_0(z_0)$ consists of two points.

4. Let Z, z_0 and $q(z)$ be the same as in the previous example. Then for the function $\Phi(z, z_0) = zz_0$ the set $\Lambda_0(z_0)$ is the closed segment $[0.5, 2]$.

In the case, when the cardinality of $\Lambda_0(z_0)$ is greater then one, we take an arbitrary element from $\Lambda_0(z_0)$ as $\lambda(z_0)$.

The function $q^0_{F,Z}(z_0) := \lambda(z_0)$ is called the *generalized polar function* to $q(z)$ on the set Z with respect to the coupling function F. By definition,

$$q^0_{F,Z}(z_0) \in \operatorname*{Argmax}_{\lambda \in \Lambda(z_0)} \Delta_{z_0}(\lambda). \tag{22}$$

Note that since $\Lambda(z_0)$ is non-empty closed segment (maybe with one of the extreme point equal to $+\infty$ or $-\infty$), the function $q^0_{F,Z}(z_0)$ is well defined. If Z coincides with the entire space $\mathbb{R}^m$, then we use notation $q^0_F(z_0)$ instead of $q^0_{F,Z}(z_0)$.

The following Minkowski–Mahler inequality follows directly from the inclusion (22):

$$F(z, z_0) \le q(z)q^0_{F,Z}(z_0), \qquad \forall\, z \in Z, \quad \forall\, z_0 \in Z_0. \tag{23}$$

If $q(z) \geq 0$ for all $z \in Z$, then the definition of $q^0_{F,Z}(z_0)$ can be rewritten in the classic form

$$q^0_{F,Z}(z_0) = \sup_{z \in Z,\ q(z)>0} \frac{F(z, z_0)}{q(z)}. \tag{24}$$

It is assumed that the supremum over empty set is equal to $-\infty$. Similarly, if $q(z) \leq 0$ for all $z \in Z$, we have:

$$q^0_{F,Z}(z_0) = \inf_{z \in Z,\ q(z)<0} \frac{F(z, z_0)}{q(z)}, \tag{25}$$

where the infimum over empty set is equal to $+\infty$.

We now present the simplest examples of polar functions.

Example 3.1. Assume that a coupling function F coincides with the classical bilinear function F_1:

$$F_1(z, z_0) = \langle z, z_0 \rangle.$$

Let $q_1(z) = \|z_+\|_p$, where $p > 1$. Let $Z = \mathbb{R}^m$. It is well known that

$$(q_1)^0_{F_1}(z_0) = \|z_0\|_{p_*},$$

where $z_0 \in Z_0 = \mathbb{R}^m_+$, $p^{-1} + p_*^{-1} = 1$.

Consider now the function $q_2(z) = -\|z_-\|_p$, where $p < 1$. Let $Z = \mathbb{R}^m_-$. Then

$$q^0_{F_1, \mathbb{R}^m_-}(z_0) = \|z_0\|_{p_*},$$

where $z_0 \in Z_0 = \mathbb{R}^m_+$. If $p \neq 0$, then $p^{-1} + p_*^{-1} = 1$. If $p = 0$, then $p_* = p = 0$.

Example 3.2. Consider now the coupling function F_2, where

$$F_2(z, z_0) = \max[\langle z, \omega_0 \rangle, \ldots, \langle z, \omega_m \rangle], \qquad z_0 = [\omega_0, \ldots, \omega_m] \in \mathbb{R}^{m(1+m)}.$$

Then we have for functions q_1 and q_2, respectively:

$$\begin{aligned} (q_1)^0_{F_2}(z_0) &= \max[\|\omega_0\|_{p_*}, \ldots, \|\omega_m\|_{p_*}], \qquad z_0 \in Z_0 = \mathbb{R}^{m(1+m)}_+, \\ (q_2)^0_{F_2, \mathbb{R}^m_-}(z_0) &= \min[\|\omega_0\|_{p_*}, \ldots, \|\omega_m\|_{p_*}], \qquad z_0 \in Z_0 = \mathbb{R}^{m(1+m)}_+. \end{aligned}$$

Example 3.3. Let F be a function defined by (16), that is, $F(z, z_0) = F_3(z, z_0)$, where

$$F_3(z, z_0) := \frac{\alpha}{2} \left[\left\| \left(\frac{z}{\alpha} + \omega_1 \right)_{[0, \omega_1 + \omega_2/\alpha]} \right\|^2 - \|\omega_1\|^2 \right],$$

and $z_0 = [\omega_1, \omega_2, \alpha] \in Z_0 = \mathbb{R}^m_+ \times \mathbb{R}^m_{++} \times \mathbb{R}_{++}$. Then we have for the function $q_1(z) = \|z_+\|_p$ (see [12]):

$$(q_1)^0_{F_3}(z_0) = \frac{1}{\|\omega_2\|_p} \left[\langle \omega_1, \omega_2 \rangle + \frac{\|\omega_2\|_2^2}{2\alpha} \right].$$

Example 3.4. Let $Z = Z_0 = \mathbb{R}^m_{++} := \{z \in \mathbb{R}^m : z^i > 0,\ \ i \in I\}$, where $I = \{1, \ldots, m\}$. Consider the coupling function $F(z, z_0) = \min_{i \in I} z^i z_0^i$. Let q be an IPH function defined on Z. Then (see [20])

$$q^0_{F, \mathbb{R}^m_{++}}(z_0) = \frac{1}{q(1/z_0)},$$

where

$$\frac{1}{z_0} = \left(\frac{1}{z_0^1}, \ldots, \frac{1}{z_0^m}\right).$$

The functions (24) and (25) possess the following simple properties:

Proposition 3.1. *Let the function $F(z, z_0)$ be convex with respect to z_0 on the convex set Z_0. Then the function* (24) *is also convex on Z_0, the function* (25) *is concave on Z_0.*

Proposition 3.2. *Suppose that $F(z, z_0)$ is an IPH function with respect to z_0. Then both functions* (24) *and* (25) *are positively homogeneous. The function* (24) *is increasing and, vice versa, the function* (25) *is decreasing.*

Proposition 3.3. *If the function $q(z)$ is positive, then $q^0_{F,Z}(z_0)$ is an increasing function of the set Z. It means that $q^0_{F,Z_1}(z_0) \leq q^0_{F,Z_2}(z_0)$ for $Z_1 \subseteq Z_2$. If $q(z)$ is negative, then $q^0_{F,Z}(z_0)$ is a decreasing function of the set Z, i.e. $q^0_{G,Z_1}(z_0) \geq q^0_{G,Z_2}(z_0)$ for $Z_1 \subseteq Z_2$.*

The proofs of Propositions 3.1 – 3.3 follow immediately from the definitions (24) and (25).

4 EXACT AUXILIARY FUNCTIONS

Let φ be an outer convolution function, ψ be an inner convolution function. Consider the corresponding separation function

$$h(u, v) = \varphi(u, \psi(v)) \tag{26}$$

and the corresponding auxiliary function

$$M(x, \eta) = \varphi(f(x) - \eta, \psi(g(x))), \tag{27}$$

defined by (3). We assume in this section that φ is a proper *increasing convolution* (IC) function. Let $Y \in \mathcal{Y}$ and $\eta \in H \subseteq \mathbb{R}$. Denote

$$X_Y(\eta) = \operatorname*{Argmin}_{x \in Y} M(x, \eta),$$

and recall that X_* is the solution set of the problem $P(f, g)$.

Definition 4.1. *The function M defined by* (27) *is said to be an exact auxiliary (EA) function on $Y \times H$, if $X_Y(\eta) = X_*$ for all $\eta \in H$.*

According to this definition the function M is an EA function on $Y \times H$ if and only if the separation function $h(u, v) = \varphi(u, \psi(v))$ enjoys the following property:

$$\operatorname*{Argmin}_{[u,v] \in \mathcal{T}_\eta(Y)} h(u, v) = \mathcal{T}_\eta(X_*) \tag{28}$$

for all $\eta \in H$. Here $\mathcal{T}_\eta(Z) = \{(f(z) - \eta, g(z)) : z \in Z\}$ is the image of a set Z under the mapping $x \mapsto (f(x) - \eta, g(x))$. It is obvious, that (28) holds, if for each $\eta \in H$ there exists a constant $\gamma = \gamma(\eta)$ such that

$$h(u, v) \geq \gamma \ \text{ for all } \ (u, v) \in \mathcal{T}_\eta(Y), \tag{29}$$

and

$$h(u, v) = \gamma \tag{30}$$

only for $(u, v) \in \mathcal{T}_\eta(X_*)$.

Below we present some sufficient conditions for M to be an EA function. These conditions depend on the problem $P(f,g)$ (more precisely, on the set $\mathcal{T}(Y)$, where $Y \in \mathcal{Y}$ and on convolution functions φ and ψ of (27).

Consider a function $F : \mathbb{R}^m \times \Omega \to \mathbb{R}$, where Ω is a set of parameters such that $F(\cdot,\omega)$ is a proper convolution function for any $\omega \in \Omega$. Let M_F be the auxiliary function, which corresponds to $\varphi(u,w) = u + w$ and $\psi(v) = F(v,\omega)$:

$$M_F(x,\omega) = f(x) + F(g(x),\omega) \qquad (x \in X,\ \omega \in \Omega).$$

Assume that the triplet (F, M_F, Y) possesses the property A (see Definition 2.1), that is, there exists a pair $(x_*, \omega_*) \in X_0 \times \Omega$ such that

$$\min_{x \in Y} M_F(x,\omega_*) = M_F(x_*,\omega_*), \qquad F(g(x_*),\omega_*) = 0. \tag{31}$$

It follows from Proposition 2.1 that x_* is a solution of $P(f,g)$, that is, $x_* \in X_*$ and $f(x_*) = \rho$.

First we assume that ψ of (27) is an exterior convolution function.

Theorem 4.1. *Let conditions (31) hold for $Y \in \mathcal{Y}$ at (x_*, ω_*). Let also φ be a proper IC function and ψ be a proper strictly exterior convolution function such that*

$$0 < \psi_F^0(\omega_*) < +\infty. \tag{32}$$

If there exists a set $H \subset \mathbb{R}$ such that the inequality

$$\varphi(u,w) > \varphi(u_*,0) \tag{33}$$

holds for all

$$u_* \in \rho - H \quad \text{and} \quad w = (u_* - u)_+/\psi_F^0(\omega_*) \quad \text{with} \quad u \neq u_*,$$

then the function M defined by (27) *is an EA on the set $Y \times H$.*

Proof. Since ψ is a strictly exterior convolution function, it follows that $h(u,v) = = h(u,0)$ for all $v \in \mathbb{R}^m_-$, where h is a separation function defined by (26). We have

$$\mathcal{T}_\eta(X_*) = \{(u,v) \in \mathbb{R}^{1+m} : u = \rho - \eta,\ \ v = g(x_*),\ x_* \in X_*\}.$$

Thus

$$h(u,v) = \varphi(\rho - \eta, 0) := \gamma(\eta) \tag{34}$$

for all $(u,v) \in \mathcal{T}_\eta(X_*)$

Let $g(Y)$ be the image of the set Y under the mapping $g : \mathbb{R}^n \to \mathbb{R}^m$. It follows from (31) and from the Minkowski–Mahler inequality (23) that for all $x \in Y$

$$\rho = f(x_*) + F(g(x_*),\omega_*) \le f(x) + F(g(x),\omega_*) \le f(x) + \psi(g(x))\psi^0_{F,g(Y)}(\omega_*). \tag{35}$$

The exterior convolution function $\psi(v)$ is nonnegative on $\mathbb{R}^m$, therefore, by Proposition 3.3,

$$\psi^0_{F,g(Y)}(\omega_*) \le \psi^0_F(\omega_*). \tag{36}$$

Combining (35) and (36), we have

$$\rho \le f(x) + \psi(g(x))\psi^0_F(\omega_*)$$

for all $x \in Y$. This inequality can be rewritten as

$$u + \psi(v)\psi^0_F(\omega_*) \ge \rho, \qquad (u,v) \in \mathcal{T}(Y). \tag{37}$$

Taking into account (32), we derive from (37) that

$$\psi(v) \geq (\rho - \eta - u)/\psi_F^0(\omega_*), \qquad (u,v) \in \mathcal{T}_\eta(Y).$$

Moreover, since $\psi(v) \geq 0$, we have

$$\psi(v) \geq (\rho - \eta - u)_+/\psi_F^0(\omega_*), \qquad (u,v) \in \mathcal{T}_\eta(Y).$$

Let $u_* = \rho - \eta$ and $w = (u_* - u)_+/\psi_F^0(\omega_*)$. The function φ is increasing on $\mathbb{R}^2$, therefore

$$h(u,v) = \varphi(u, \psi(v)) \geq \varphi(u,w).$$

If $u \neq u_*$, then $(u,v) \notin \mathcal{T}_\eta(X_*)$. Moreover, due to (33),

$$\varphi(u,w) > \varphi(u_*, 0) = \varphi(\rho - \eta, 0) := \gamma(\eta). \tag{38}$$

Therefore, $h(u,v) > \gamma(\eta)$.

Suppose now that $u = u_*$ and $(u,v) \notin \mathcal{T}_\eta(X_*)$. Then $v \notin \mathbb{R}^m_-$, that is, $\psi(v) > 0$. Since φ is increasing, we have

$$h(u,v) = \varphi(u_*, \psi(v)) \geq \varphi(u_* - \psi_F^0(\omega_*)\psi(v), \psi(v)). \tag{39}$$

Let $u_1 = u_* - \psi_F^0(\omega_*)\psi(v)$. For such u_1 we have $u_1 \neq u_*$ and $w = (u_* - u_1)_+/\psi_F^0(\omega_*) = \psi(v)$. Thus, due to (33), we obtain

$$\varphi(u_* - \psi_F^0(\omega_*)\psi(v), \psi(v)) = \varphi(u_1, w) > \varphi(u_*, 0) = \gamma(\eta).$$

It follows from (34), (38) and (39) that both (29) and (30) hold. Therefore, $M(x,\eta)$ EA function on $Y \times H$ $\square$

The inequality (33) has a simple geometrical interpretation. Denote by $K_1(u_*)$ the shift along the horizontal axis of the cone

$$K_1 = \{(u,w) \in \mathbb{R}^2 : w \geq (-u)_+/\psi_F^0(\omega_*)\},$$

that is, $K_1(u_*) = (u_*, 0) + K_1$. According to (33), the function $\varphi(u,v)$ enjoys the following property: for any $u_* = \rho - \eta$ with $\eta \in H$ the level line of $\{(u,v) : \varphi(u,v) = \varphi(u_*, 0)\}$ of the function φ does not intersect the set $K_1(u_*)$ with the exception of the point $(u_*, 0)$.

We now present three examples of EA functions.

Example 4.1. Let $\alpha < 0$. Then the function

$$M(x,\eta) = \begin{cases} (\eta - f(x))_+^\alpha + \psi(g(x)), & f(x) < \eta, \\ +\infty, & f(x) \geq \eta, \end{cases} \tag{40}$$

is EA function on $\mathbb{R}^n \times H$, where

$$H = \{\eta \in \mathbb{R} : \eta > \rho + (\alpha\psi_F^0(\omega_*))^{1/(1-\alpha)}\}. \tag{41}$$

Example 4.2. Let $\beta > 1$. Then the function

$$M(x,\eta) = (f(x) - \eta)_+^\beta + \psi(g(x)) \tag{42}$$

is EA function on $\mathbb{R}^n \times H$, where

$$H = \{\eta \in \mathbb{R} : \rho \geq \eta > \rho - (\beta\psi_F^0(\omega_*))^{1/(1-\beta)}\}. \tag{43}$$

Example 4.3. Let $t = 1 - (f(x) - \eta)_{-}\psi(g(x))$t. The function

$$M(x,\eta) = \begin{cases} \dfrac{(f(x)-\eta)}{t}, & t > 0, \\ -\infty, & t \leq 0, \end{cases} \tag{44}$$

is EA on the set $\mathbb{R}^n \times H$, where

$$H = \left\{\eta \in \mathbb{R} : \eta > \rho + (\psi_F^0(\omega_*))^{1/2}\right\}. \tag{45}$$

The functions (40), (42) and (44) were considered in [10]]. However, the lower and upper boundaries for η providing the exactness of auxiliary functions were obtained only under assumption that there exists the saddle point of Lagrange function, and thus the usual polar function was used in (41), (43) and (45) instead of generalized polar function. As it has been shown in [12], for functions (40) and (42) this assumption may be replaced by the sufficient conditions given by Proposition 2.1.

We now assume that the function ψ in (27) is a proper interior convolution function and $Y = X_0$.

Theorem 4.2. *Let $Y = X_0$ and conditions (31) be satisfied at (x_*, ω_*). Let also φ be a proper IC function and ψ be a proper interior convolution function such that*

$$0 < \psi^0_{F,\mathbb{R}^m_-}(\omega_*) < +\infty. \tag{46}$$

If there exists a set $H \subset \mathbb{R}$ such that the inequality

$$\varphi(u,w) > \varphi(u_*,0) \tag{47}$$

holds for all $u_* \in \rho - H$ and $w = (u_* - u)/\psi^0_{F,\mathbb{R}^m_-}(\omega_*)$ with $u > u_*$, then the function $M(x,\eta)$ is EA on the set $X_0 \times H$.

Proof. Since $g(x) \in \mathbb{R}^m_-$ and $\psi(g(x)) \leq 0$ for all $x \in X_0$ we have, combining (31), (23) and Proposition 3.3,

$$\begin{aligned} \rho = f(x_*) + F(g(x_*),\omega_*) &\leq f(x) + F(g(x),\omega_*) \leq f(x) + \psi(g(x))\psi^0_{F,g(X_0)}(\omega_*) \leq \\ &\leq f(x) + \psi(g(x))\psi^0_{F,\mathbb{R}^m_-}(\omega_*) \leq f(x). \end{aligned}$$

We derive from these inequalities that

$$(\rho - f(x))/\psi^0_{F,\mathbb{R}^m_-}(\omega_*) \leq \psi(g(x)) \leq 0 \tag{48}$$

for all $x \in X_0$.

Let $u = f(x) - \eta$, $u_* = \rho - \eta$ and $w = (u_* - u)/\psi^0_{F,\mathbb{R}^m_-}(\omega_*)$. Then the inequalities (48) can be rewritten as

$$w \leq \psi(v) \leq 0 \quad \text{for all} \quad (u,v) \in \mathcal{T}_\eta(X_0). \tag{49}$$

Let $\eta \in H$ and $(u,v) \in \mathcal{T}_\eta(X_0)$. If $u \neq u_*$, then, combining (49) and (47), we have

$$h(u,v) = \varphi(u, \psi(v)) \geq \varphi(u,w) > \varphi(u_*, 0) := \gamma(\eta).$$

If $u = u_*$, then $(u,v) \in \mathcal{T}_\eta(X_*)$ and $w = 0$. ue to (49), we have $\psi(v) = 0$. Therefore,

$$h(u,v) = \varphi(u_*, 0) = \gamma(\eta).$$

Thus, $M(x,\eta)$ is EA function on the set $X_0 \times H$. $\square$

The inequality (47) means geometrically that for any $u_* = \rho - \eta$ with $\eta \in H$ the level line $\{(u,w) : \varphi(u,w) = \varphi(u_*, 0)\}$ of the function φ does not intersect the set

$$K_2(u_*) = \left\{(u,w) \in \mathbb{R}^2 : w \geq (u_* - u)/\psi^0_{F,\mathbb{R}^m_-}(\omega_*), \;\; u \geq u_*\right\},$$

excepting the point $(u_*, 0)$.

The functions (40), (42) and (44) with an interior convolution function ψ satisfy conditions of Theorem 4.2. All these functions are EA on the set $X_0 \times H$, where the set H is equal to, respectively:

$$\begin{aligned} H &= \left\{\eta \in \mathbb{R} : \rho < \eta < \rho + (\alpha \psi^0_{F,\mathbb{R}^m_-}(\omega_*))^{1/(1-\alpha)}\right\}, \\ H &= \left\{\eta \in \mathbb{R} : \eta < \rho - (\beta \psi^0_{F,\mathbb{R}^m_-}(\omega_*))^{1/(1-\beta)}\right\}, \\ H &= \left\{\eta \in \mathbb{R} : \rho \leq \eta < \rho + \left(\psi^0_{F,\mathbb{R}^m_-}(\omega_*)\right)^{1/2}\right\}. \end{aligned}$$

Assume now that $\psi(v)$ is a strictly common convolution function.

Theorem 4.3. *Lef $Y \in \mathcal{Y}$ and conditions* (31) *be satisfied at (x_*, ω_*). Let also φ be a proper IC function and ψ be a strictly common convolution function such that* (32) *and* (46) *hold. Assume that there exists a set $H \subset \mathbb{R}$ such that*

$$\varphi(u_*, 0) < \begin{cases} \varphi(u, (u_* - u)/\psi^0_F(\omega_*)), & u < u_*, \\ \varphi(u, (u_* - u)/\psi^0_{F,\mathbb{R}^m_-}(\omega_*)), & u > u_*. \end{cases}$$

for all $u_* \in \rho - H$ and $u \neq u_*$. Then $M(x,\eta)$ is an EA function on the set $Y \times H$.

The proof is similar to those of Theorems 4.1 and 4.2.

Let ψ be a strictly common convolution function such that $\psi^0_F(\omega_*) < \psi^0_{F,\mathbb{R}^m_-}(\omega_*)$. Then the function (44) is EA on the set $Y \times H$, where the set H is equal to

$$H = \left\{\eta \in \mathbb{R} : \rho + \left(\psi^0_F(\omega_*)\right)^{1/2} < \eta < \rho + \left(\psi^0_{F,\mathbb{R}^m_-}(\omega_*)\right)^{1/2}\right\}.$$

A more general than (27) auxiliary function

$$M(x,\omega) = \varphi(\phi(f(x),\omega), \psi(g(x))), \tag{50}$$

was also considered in [10], where $\phi(u,\omega)$ is an increasing convex function with respect to $u \in R$.

Let $\omega \in \mathbb{R}$. If $\phi(u,\omega) = u - \omega$, then the function (50) coincides with (27). Let now $\phi(u,\omega) = u/\omega$. Assume that $\rho > 0$ and ψ is a strictly exterior convolution function. Assume also that the function φ and the set $H \subset \mathbb{R}$ are such that the inequality (33) holds for all $u_* \in H$, for all $u \neq u_*$, and for $w = (u_* - u)_+/\psi^0_F(\omega_*)$. Then the function

$$M(x, \omega) = \varphi(f(x)/\omega, \psi(g(x)))$$

is EA on the set $Y \times \Omega_*$, where $\Omega_* = \{\omega \geq 1 : \rho/\omega \in H\}$. Similar results may be obtained for the general auxiliary function (50). Moreover, the function ψ may be of any type (strictly exterior, interior or strictly common convolution function). Note also that it is possible to consider the following generalization of the function (50)

$$M(x, \eta, \omega) = \varphi(\phi(f(x) - \eta, \omega), \psi(g(x))).$$

For such a function the interval H of values η providing the exactness of the auxiliary function may be essentially extended by choosing an appropriate value of the parameter ω.

Remark 4.1. *All auxiliary functions* (40), (42) *and* (44) *may be used for solving constrained optimization problem* $P(f, g)$. *If* $\psi^0_{F,g(Y)}(\omega_*)$ *is a positive finite number, then we have an opportunity to find a global solution of* $P(f, g)$ *by a single minimization of an unconstrained function.*

Acknowledgements

The authors are very thankful to anonymous referees for their valuable comments. This Research has been supported by Australian Research Council (Grant No. A49906152) and by Russian Foundation for Basic Research (Grants No. 99-01-01186 and No. 00-15-96080).

REFERENCES

[1] M.Yu. Andramonov (2000). An approach to constructing generalized penalty function. In: A.M. Rubinov and B.M. Glover (Eds.). *Optimization and related topics*, 3–14. Kluwer Academic Publishers, Dordrecht.

[2] M.S. Bazaraa and J.J. Goode (1982). Sufficient conditions for a globally exact penalty function without convexity. *Mathematical Programming Study*, **19**, 1–15.

[3] D.P. Bertsecas (1982). *Constrained Optimization and Lagrange Multiplier Methods*, Academic Press, New-York.

[4] J. Burke (1991). An exact penalization viewpoint of constrained optimization. *SIAM J. Control and Optimization*, **29**, 968–998.

[5] Ch. Charalambous (1976). A negative-positive barrier method for non-linear programming, *Internat. J. Systems Sci.*, **7**, 557–575.

[6] V.F. Demyanov, G. Di Pillo and F. Facchinei (1998). Exact penalization via Dini and Hadamard conditional derivatives. *Optimization Methods and Software*, **9**, 19–36.

[7] S. Dolecky and S. Rolewich (1979). Exact penalties for local minima. *SIAM J. Control and Optimization*, **17**, 596–606.

[8] I.I. Eremin (1967). The "penalty" method in convex programming. *Dokl. Akad. Nauk SSSR*, **173**, 748–751 (in Russian).

[9] Yu.G. Evtushenko (1987). Accuracy estimation for penalty function methods. In: *Problems of Applied Mathematics and Computer Science*, 199–208, Nauka, Moscow (in Russian).

[10] Y.G. Evtushenko and V.G. Zhadan (1990). Exact auxiliary functions in optimization problems. *U.S.S.R. Comput. Maths. Math. Phys.*, **30**, 31–42.

[11] Y.G. Evtushenko and V.G. Zhadan (1990). Exact auxiliary functions, *Informatica*, **1**, 40–58.

[12] Y.G. Evtushenko and V.G. Zhadan (1992). Exact auxiliary functions in non-convex optimization. In: W. Oettli and D. Pallaschke (Eds.), *Advances in Optimization*, (Lecture Notes in Economics and Mathematical Systems, 382), 217–226, Springer-Verlag, Berlin.

[13] Y.G. Evtushenko, A.M. Rubinov and V.G. Zhadan (2001). General Lagrange-type functions in constrained global optimization. Part I: Auxiliary functions and Optimality Conditions. *Optimization Methods and Software*, **16**, 179–217.

[14] S.P. Han and O.L. Mangasarian (1979). Exact penalty functions in nonlinear programming. *Mathematical Programming*, **17**, 251–269.

[15] P. Huard (1964). Resolution des P.M. a constraintes non-lineaires par la methode des centres, Note E.D.F. HR 5.690, May 6.

[16] J. Kowalik, M.R. Osborn and D.M. Ryan (1969). A new method for constrained optimization problems. *Operat. Res.*, **17**, 973–989.

[17] D.D. Morrison (1968). Optimization by least squares. *SIAM J. Numer. Analysis*, **5**, 83–88.

[18] G. Di Pillo and L. Grippo (1989). Exact penalty functions in constrained optimization. *SIAM J. Control Optim.*, **27**, 1333–1360.

[19] D. Pollaschke and S. Roleviwics (1997). *Foundations of Mathematical Optimization.* Kluwer Academic Publishers.

[20] A.M. Rubinov (2000). *Abstract Convexity and Global Optimization.* Kluwer Academic Publishers, Dordrecht.

[21] A.M. Rubinov, B.M. Glover and X.Q. Yang (1999). Extended Lagrange and penalty functions in continuous optimization. *Optimization*, **46**, 327–351.

[22] A.M. Rubinov, B.M. Glover and X.Q. Yang (2000). Decreasing functions with application to penalization. *SIAM J. Optim.*, **10**, 289–313.

[23] A.M. Rubinov and A. Uderzo. On global optimality conditions via separation functions. *Journal of Optimization Theory and Applications* (to appear).

[24] I. Singer (1997). *Abstract Convex Analysis.* Wiley and Sons.

[25] V.V. Velichenko (1967). A method for determining the constrained minimum of a function of many variables. *Avtomatika i telemechan.*, No.2, 171–172 (in Russian).

[26] X.Q. Yang and K.L. Teo. Nonlinear Lagrangian functions and applications to semi-infinite programs, optimization and numerical algebra. *Annals of Operations Research* (to appear).

[27] W. Zangwill (1967). Nonlinear programming via penalty function. *Management Sciences*, **13**, 344–358.

[28] V.G. Zhadan (1984). On some estimations of penalty coefficient in methods of exact penalty functions. *U.S.S.R. Comput. Maths. Math. Phys.*, **24**, 124–128.